(2)

IDA DOCUMENT D-840

ISSUES IN MEASURING COST GROWTH

Karen W. Tyson, *Project Leader*
Stephen J. Balut
Neang I. Om
Stephen K. Welman

September 1990

INSTITUTE FOR DEFENSE ANALYSES
1801 N. Beauregard Street, Alexandria, Virginia 22311-1772

IDA Log No. HQ 90-36109

DEFINITIONS

IDA publishes the following documents to report the results of its work.

Reports

Reports are the most authoritative and most carefully considered products IDA publishes. They normally embody results of major projects which (a) have a direct bearing on decisions affecting major programs, (b) address issues of significant concern to the Executive Branch, the Congress and/or the public, or (c) address issues that have significant economic implications. IDA Reports are reviewed by outside panels of experts to ensure their high quality and relevance to the problems studied, and they are released by the President of IDA.

Group Reports

Group Reports record the findings and results of IDA established working groups and panels composed of senior individuals addressing major issues which otherwise would be the subject of an IDA Report. IDA Group Reports are reviewed by the senior individuals responsible for the project and others as selected by IDA to ensure their high quality and relevance to the problems studied, and are released by the President of IDA.

Papers

Papers, also authoritative and carefully considered products of IDA, address studies that are narrower in scope than those covered in Reports. IDA Papers are reviewed to ensure that thev meet the high standards expected of refereed papers in professional journals or formal Agency reports.

Documents

IDA Documents are used for the convenience of the sponsors or the analysts (a) to record substantive work done in quick reaction studies, (b) to record the proceedings of conferences and meetings, (c) to make available preliminary and tentative results of analyses, (d) to record data developed in the course of an investigation, or (e) to forward information that is essentially unanalyzed and unevaluated. The review of IDA Documents is suited to their content and intended use.

The work reported in this publication was conducted under IDA's Independent Research Program. Its publication does not imply endorsement by the Department of Defense, or any other Government agency, nor should the contents be construed as reflecting the official position of any Government agency.

UNCLASSIFIED

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1990	Final Report, Dec 89 – Sep 1990

4. TITLE AND SUBTITLE

Issues in Measuring Cost Growth

5. FUNDING NUMBERS

IDA Independent Research Program

6. AUTHOR(S)

Karen W. Tyson, Stephen J. Balut, Neang I. Om, and Stephen K. Welman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311-1772

8. PERFORMING ORGANIZATION REPORT NUMBER

IDA-D-840

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Not applicable.

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12A. DISTRIBUTION/AVAILABILITY STATEMENT

Suitable for public release.

12B. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The authors examine measures of cost growth for major weapon systems. Measures used in the past for both single and multiple programs and for both defense and and non-defense systems are reviewed. The advantages and disadvantages of these measures are discussed and recommendations for appropriately measuring and forecasting cost growth in the future are presented.

14. SUBJECT TERMS

Cost Overruns, Weapon Systems, Acquisition, Cost Models

15. NUMBER OF PAGES

34

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

UNCLASSIFIED

IDA DOCUMENT D-840

ISSUES IN MEASURING COST GROWTH

Karen W. Tyson, *Project Leader*
Stephen J. Balut
Neang I. Om
Stephen K. Welman

September 1990

IDA

INSTITUTE FOR DEFENSE ANALYSES

IDA Independent Research Program

PREFACE

This document was prepared by the Institute for Defense Analyses (IDA) under IDA's Independent Research Program. The objective of the work was to discuss ways of measuring cost growth in the acquisition of major defense systems.

The work presented in this document was reviewed by the Cost Analysis and Research Division (CARD).

CONTENTS

FIGURES

TABLE

I. INTRODUCTION

Public perception is that cost estimates for defense items are wildly inaccurate and that defense contracts are chronically plagued by overruns. As part of an effort to document and contain cost overruns in major weapon systems, the Department of Defense (DoD) and others began to develop measures of "cost growth," the difference between the planned cost of the system and the actual cost.

Among the reasons why a good measure of cost growth is necessary are the following:

- Individual program management. The Office of the Secretary of Defense (OSD) and the services want to structure programs for effective accomplishment of objectives. OSD can use cost growth measures to monitor the progress of existing programs for Defense Acquisition Board and other reviews. OSD will want to encourage and continue programs with low cost growth and to emphasize cost containment efforts on programs with high cost growth.
- Comparison of programs. A good cost growth measure would allow for evaluation of competing programs.
- Congressional monitoring. Congress wants to be able to determine how well a program is proceeding, and a good cost growth measure would be useful.
- Assessment of acquisition program management. The Comptroller regularly issues the Selected Acquisition Report (SAR) summary tables, which serve as a yardstick of how all programs are doing. Several studies have involved the use of cost growth measures as a starting point to evaluate overall acquisition experience. The results of these studies can then be used to evaluate the effectiveness of cost estimating, changes in acquisition strategy, and management initiatives in improving outcomes. Such measures are needed to assess, for example, the effectiveness of the Packard initiatives of the early 1970s or to provide a baseline for evaluating 1980s initiatives, such as fixed-price development and dual sourcing.

The objective of this document was to examine measures of cost growth in major systems acquisition. Our approach was to examine the literature on cost growth and to assess the advantages and disadvantages of the measures used. We then compiled our thoughts on what could be done to improve existing measures of cost growth and to develop new measures.

II. PAST MEASURES OF COST GROWTH

The Selected Acquisition Reports (SARs) are issued quarterly by the DoD and provide information that can be used to develop measures of cost growth. The SARs are status reports from acquisition program managers that contain the latest estimates of progress in achieving technical, schedule, and cost goals.

The baseline is typically the estimates of technical, schedule, and cost goals at the time that the program entered full-scale development (FSD). Technical goals vary according to the type of equipment, and cover the weapon system's performance and technical characteristics. The schedule goals reported typically include dates for FSD contract award, initial operational capability (IOC), and the various acquisition milestones. The cost goals include costs for development, production, military construction, and other program costs in both current and constant dollars. Each program has an established base year, typically the FSD year.

As the program progresses, variances from planned costs are reported in the following categories:

- Economic
- Quantity
- Schedule
- Engineering
- Estimating
- Support
- Other.

The SAR information is often used as a starting point for measuring cost growth. In the remainder of this section, we discuss various methods doing so. First, we discuss methods used for measuring cost growth of a single weapons program. Then, we explore measures used for multiple programs. Finally, we report on methods employed for non-defense programs.

A. SINGLE-PROGRAM MEASURES

The most naive measure of cost growth is a simple ratio of the difference between the current estimate (CE) and the development estimate (DE), divided by the development estimate:

$$\text{Cost Growth} = (CE - DE)/DE \ .$$

See, for example, Reference [1].

1. Adjusting for Quantity and Escalation

One problem that occurs in evaluating cost growth is that program plans change. The quantity to be acquired may increase or decrease. Cost analysts make estimates based on a particular quantity, and it does not seem reasonable to define increases in cost due to increases in quantity as cost growth.

Adjustments to estimates to account for changes in quantity might be calculated in one of two ways. The first way adjusts the development estimate up to the current estimate quantity:

$$CG = (CE - (DE + Q))/(DE + Q) \ ,$$

where Q is the quantity variance. The second way adjusts the current estimate down to the development estimate:

$$[(CE - Q) - DE]/DE \ .$$

The first measure tells us what the development estimate would have been for the quantity. The second measure tells us, what the current estimate would be for the originally planned quantity. The first measure may be useful in evaluation of a single program; however, it has the problem of creating a "floating baseline" in that the baseline changes as program quantity requirements change [2]. Adjusting to the development estimate quantity, the second option, has been used by most evaluators interested in documenting program histories or in evaluating multiple programs [1 through 9]. In that case, each program is evaluated against a fixed baseline.

There are different methods for determining Q, the quantity variance. The SAR gives the program manager's best estimate of the quantity variance. Many evaluators have used this estimate [1 through 7]. The DoD Instruction for determining this variance requires use of the cost/quantity curve developed at FSD start [10]. However, when working with historical SARs, it is not clear that this guidance is widely known or

followed. In addition, other quantity-related variance is contained in other variance categories.

An alternative measure has been developed that involves independent estimation of the price-improvement curve. The actual price-improvement curve is estimated, and the point at which the development estimate quantity falls is called the DE [8, 9]. We believe that this uniform adjustment method makes more sense.

Adjusting for escalation—e.g., working in constant "base year" dollars—is important in measuring program management effects. In this formulation, cost growth due to economic escalation (changes in anticipated inflation rates) is removed. In some formulations, cost growth due to "program change escalation" is also removed. Program change escalation is the additional escalation due to program changes [1]. For example, a program "stretch" that involves acquiring items in later years would add to accrued inflation.

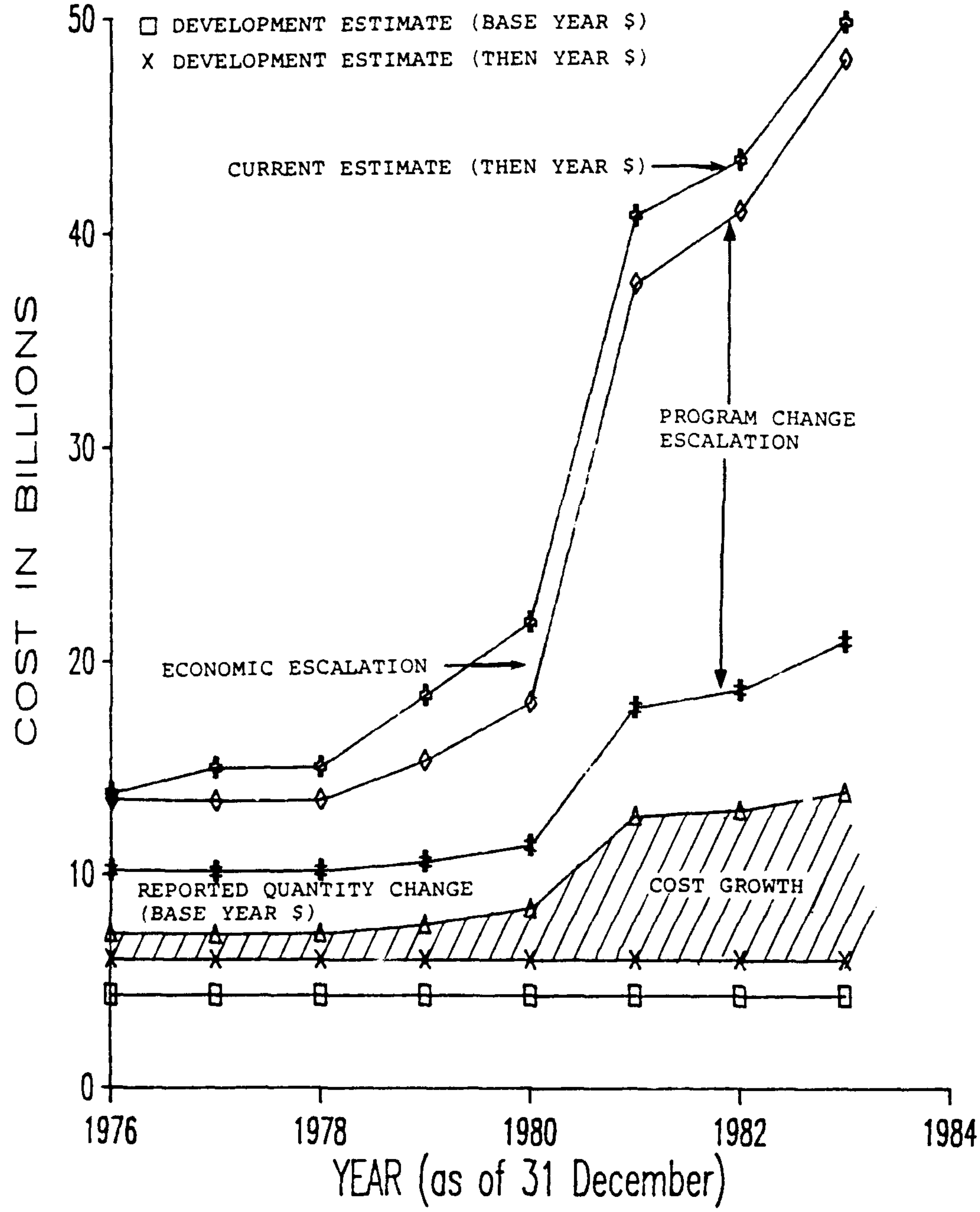
$$\text{Cost Growth} = [(\text{CE} - \text{ESC CE}) - \text{DE} + \text{Q} - \text{ESC DE})]/(\text{DE} + \text{Q} - \text{ESC DE}).$$

Figure 1 shows an example of this using the F-16 program. The various types of cost growth have been decomposed. The figure indicates that measures of cost growth can run from 70 percent to 700 percent, depending on the definition [1].

2. RAND's Work

Among a series of RAND studies on weapons acquisition, the one on acquisition policy effectiveness in the 1970s [2] was particularly detailed in terms of methodology. The single-program measures used include the ratio of the CE to the DE, in constant dollars, adjusting quantity back to the DE quantity. This measure proved to be a vast improvement over earlier measures. Working in terms of the development estimate quantity avoids the problem of a floating baseline.

The RAND measure did have some problems, however. The report discusses a correct quantity adjustment measure using the learning curve. However, this adjustment was made by subtracting the quantity variance given in the SAR from the CE. As previously noted, the methods used by program offices for calculating quantity variance may not be uniform, and we prefer an objective, replicable measure using the price-improvement curve. The study also failed to separate development and production cost growth.



Source: Reference [1].

Figure 1. F-16 Program Cost History

In addition to measuring cost growth at a point in time, RAND plotted cost growth ratios for a program at each year after Milestone II. These ratios indicated that many programs exhibit very little cost growth in their early stages, but that there is a general tendency for cost growth to increase over time. This increase was attributed to a number of factors; including:

- Unforeseen technical problems
- Changes in scope and system performance
- Schedule slippage to meet funding or performance requirements
- Changes in estimating due to better information.

Unlike the Comptroller, the RAND studies omit planning and planning/ development estimates in calculating cost growth. They include only those programs that have a development estimate.

3. IDA's Work

An IDA study on effective acquisition initiatives [9] provided a measure of individual program cost growth that offered several advantages over past measures. (Schedule outcomes were also presented.) First, unlike the RAND studies, IDA separated development and production cost growth. This allowed for separate analyses of these phenomena, which have different patterns and causes.

Second, IDA performed the quantity adjustment in a uniform way across programs. All quantities were adjusted back to the development estimate quantity. This allowed for a fixed baseline of the program. Instead of using the quantity variance given in the SAR, the price-improvement curve on actual experience was used to adjust for quantity change. This approach provided an objective measure for comparing program outcomes.

4. MCR's Work

Work conducted at Management Consulting and Research (MCR) reported on variance analysis and performance indices and factors [1, 3, 4], but these are more oriented to the contract level than the program level. For example, cost variance can be measured by comparing the budgeted cost of work performed with the actual cost of work performed.

In the Defense Acquisition Executive Summary (DAES) Analyst's Guide, MCR used two methods for cost growth measurement: (1) Variance Analysis and (2) Performance Indices and Factors. The terms and concepts were defined and described at the contract cost level and, where appropriate, extended first to the contract price level, and

finally to the program level, where such terms were applicable. The terms used in the analysis were:

- BCWS (Budgeted Cost of Work Scheduled)
- BCWP (Budgeted Cost of Work Performed)
- ACWP (Actual Cost of Work Performed)
- BPWS (Budgeted Price of Work Scheduled)
- BPWP (Budgeted Price of Work Performed)
- APWP (Actual Price of Work Performed).

BCWS is the dollar amount budgeted for work to be accomplished by a given time. This time-phased budget plan represents what the contractor is trying to achieve, and cost and schedule performance are measured against it. (It is used as the baseline.)

BCWP is the dollar amount budgeted for work actually accomplished by a given time. It represents the earned value of work performed in terms of the original budget (BCWS). BCWP is plotted only up to the "as of" date of the last cost performance report, while BCWS can be plotted out to the end of the contract, since it represents a plan. When compared to the original plan, BCWP shows how many dollars are to be credited for the amount of work performed through the "as of" date. A BCWP below the BCWS means that the project is behind schedule. The MCR report gives no indication of how to measure the work performed.

ACWP is the cost actually incurred by the contractor by a given time. This measure is usually described in terms of cumulative expenditures to date and is represented graphically along with BCWS (the plan) and BCWP (the amount budgeted for work accomplished). ACWP can only be plotted up to the "as of" date in the latest cost performance report. An ACWP above the BCWP means more dollars have been spent on work accomplished than planned.

At the program level, BPWP is the program manager's best estimate of that portion of the BPWS that actually was accomplished. It is the earned value for work performed against a specific appropriation shown in the SAR. APWP (program level) is the program manager's best estimate of the total price paid for all goods and services (work performed) recorded against a specific appropriation shown in the SAR.

Cost Variance (CV) is a measure of how closely actual expenditures are tracking to planned expenditures for work performed. (See Figure 2.) It can be measured in dollars or as a percentage:

$$CV = BCWP - ACWP$$

or

$$CV\% = (CV/BCWP) * 100$$

When CV < 0, more dollars have been spent for a given amount of work than planned. When CV% < 0, costs are overrun.

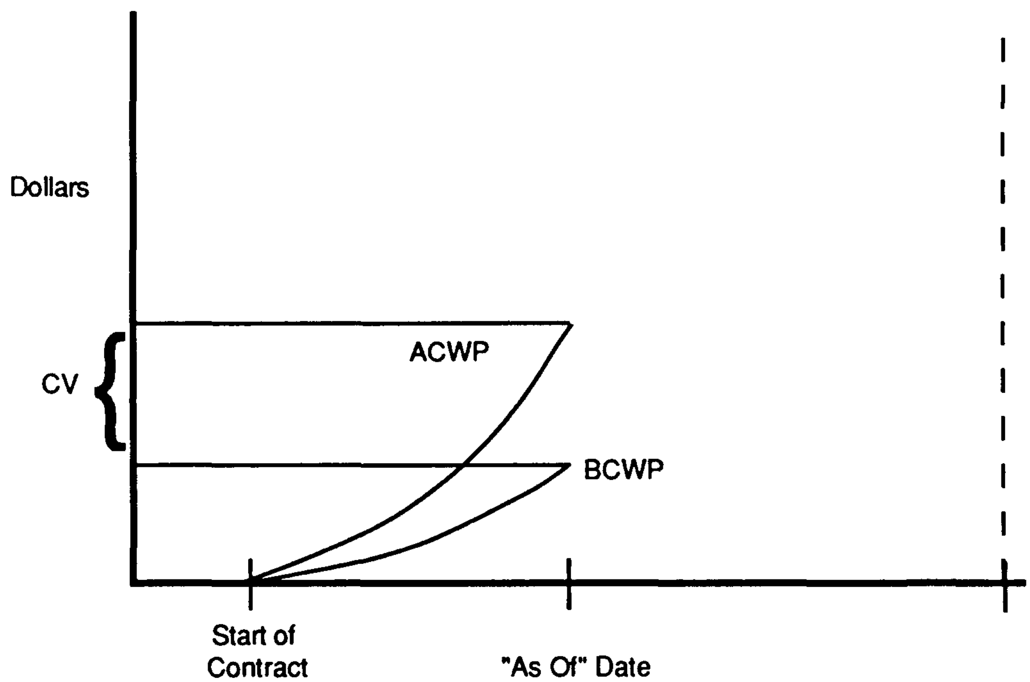


Figure 2. Cost Variance

Schedule Variance (SV) is a measure of how closely work is progressing according to plan. (See Figure 3.) It can be measured in dollars or as a percentage of work scheduled:

$$SV = BCWP - BCWS$$

or

$$SV\% = (SV/BCWS) * 100$$

When SV < 0, the project is behind schedule. In this sense, SV can be measured in both dollars and time. If SV% < 0, then the project is behind schedule.

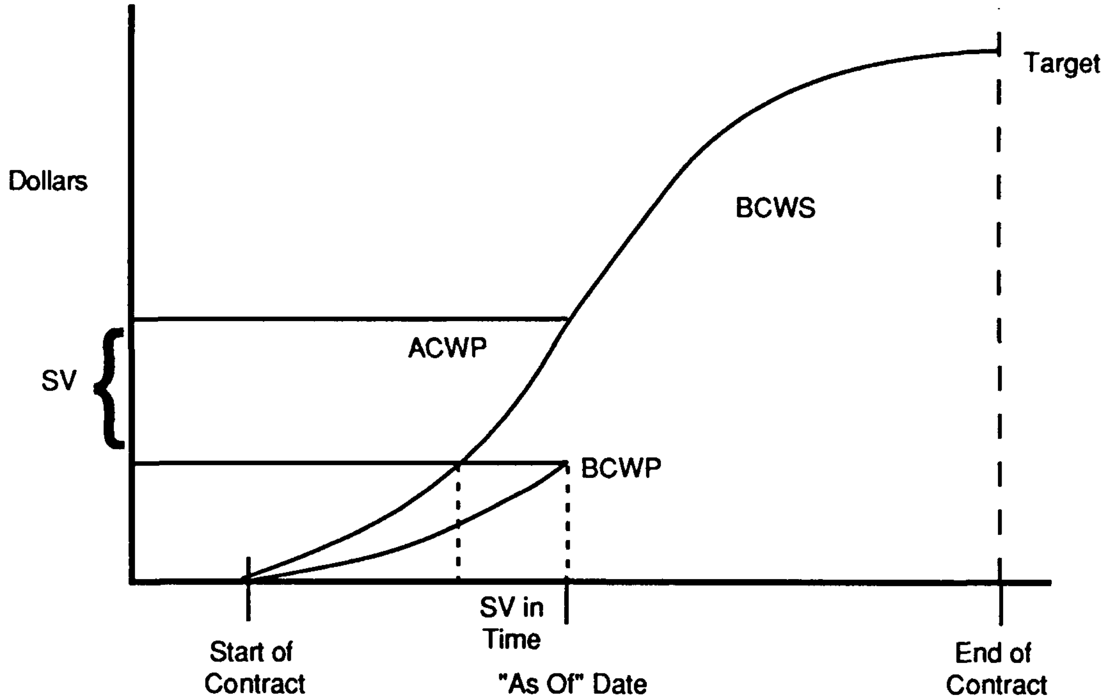


Figure 3. Schedule Variance

The Cost Performance Index (CPI) provides a measure of cost efficiency with which work has been accomplished. The formula is:

$$CPI = BCWP/ACWP$$

If CPI = 1, the program is progressing according to budget, but if CPI < 1, actual costs exceeded budgeted costs.

The Schedule Performance Index (SPI) provides a measure of the schedule efficiency with which work has been accomplished:

$$SPI = BCWP/BCWS$$

Percent Complete (%Complete) is a measure of the status of contract work in relation to the overall budget. It represents the relationship between actual progress (BCWP) and the total budgeted dollar value of the contract, the budget at completion (BAC). The formula is:

$$\%Complete = BCWP/BAC$$

Percent Spent (%Spent) is the portion of the budget spent to date. It is the ratio of actual expenditure (ACWP) and the budget at completion (BAC):

$$\%Spent = ACWP/BAC$$

The "To Complete" Performance Index (TCPI) measures the amount of "work left" with the amount "money left."

$$TCPI = Work\ Left/Money\ Left = (BAC - BCWP)/(BAC - ACWP)$$

If TCPI = 1, the program is running according to plan, but if TCPI >1, "work left" exceeds "money left."

The strength of this approach is its comprehensiveness. Weaknesses include:

- It is more oriented to contract level
- There are no specific milestones or delivery dates for work performed
- It provides for no quantity adjustment.

B. MULTIPLE-PROGRAM MEASURES

1. The "Hump Chart"

The Comptroller issues a report each year with the SAR summary tables. This report gives two cost growth measures: percentage of cost growth to date in quantity-adjusted base-year dollars, and percentage of cost growth to date in quantity-adjusted then-year dollars. Quantity adjustment appears to be done according to the SAR variance categories, adjusting the development estimate to the current estimate quantity. This procedure has some advantage if one is interested in monitoring individual programs, because it brings the dollar magnitudes closer to actuals. However, it has the problem of creating a floating baseline.

A total cost variance is also provided for the current quarter and to date. The total cost variance is broken down into the standard SAR cost variance categories.

Baselines depend on the stage of the program, and are referred to as planning estimate (PE, calculated at Milestone I), development estimate (DE, calculated at Milestone II), and production estimate (PdE, calculated at Milestone III). The production estimate is a relatively new measure. In the early 1980s, programs were given the option of re-baselining at Milestone III. In 1985, this re-baselining became mandatory.

The Comptroller also maintains a measure of annual rates of program cost growth for selected SAR weapon systems, the "hump chart" (see Figure 4). This chart is developed using a method discussed in a special study by the Congressional Budget Office (CBO) [7].

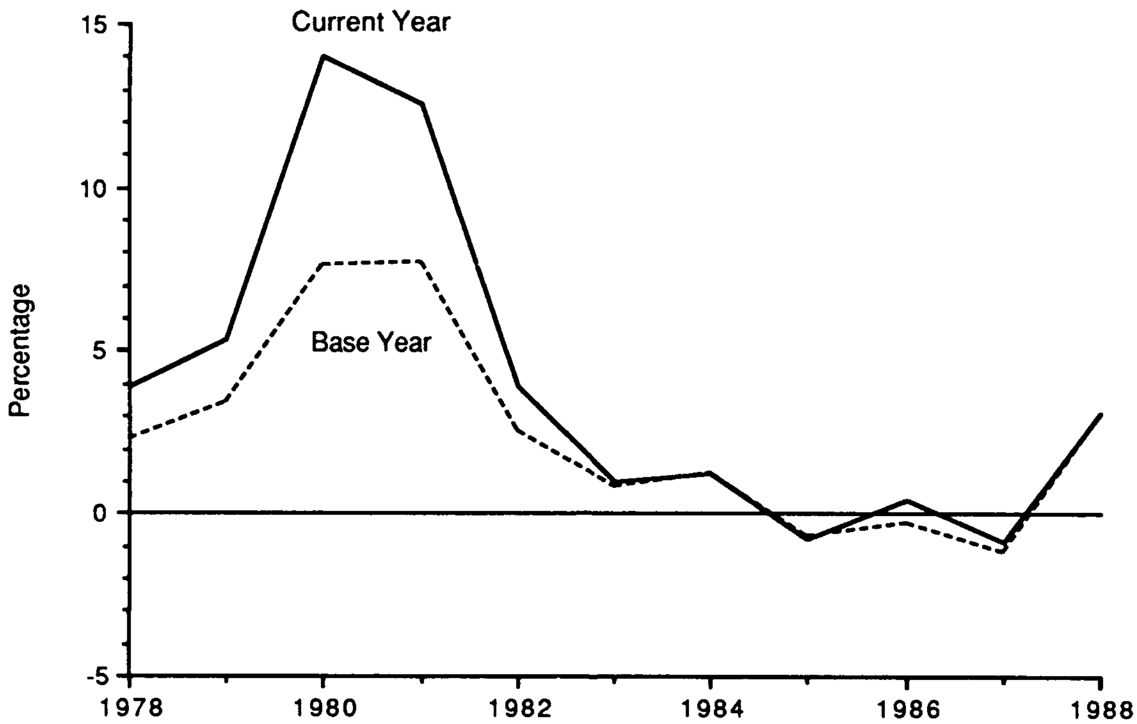


Figure 4. Annual Rates of Program Cost Growth for Selected SAR Weapon Systems (Excluding Economic and Quantity Changes)

For each year, the new cost growth posted to the SARs is evaluated. All new non-quantity cost variances (using different base years) are added up and divided by the total development estimate (again using different base years). The mixing of base years can distort cost growth experience. For example, base-year dollars might mean 1982 dollars for one program and 1975 dollars for another.

In addition, the cost growth represented here includes both past cost growth and anticipated cost growth. The accuracy of projected cost growth will not be known until all of the systems have been produced and delivered.

During the early 1980s, some policymakers speculated that cost growth was a thing of the past. This perception was based on the relatively small cost growth during that time,

as shown in the hump chart. In the late 1980s, cost growth as measured by the hump chart increased. Several factors might distort these early 1980s estimates:

- New Programs. During the 1980s, there was an unusually large number of new starts. In the 1981-85 period alone, at least 18 major programs went into FSD. As shown in a number of studies, cost growth is relatively low in new programs, and tends to increase as uncertainty decreases and reality is faced.
- Program Funding. Programs in the 1980s were amply funded, due to the defense buildup. Therefore, cost analysts had less incentive to be optimistic about costs when doing the development estimate.
- Inflation Dividend. During the late 1970s, inflation outran program funding requirements. During the early 1980s, the rate of inflation declined, and the DoD (and many other forecasters) overestimated inflation. This resulted in program overfunding in then-year dollar terms. Again, this made it seem like there was less cost growth.
- Programs Not Included. During the 1980s, there were several major "black" programs, particularly those involving Stealth technology. These programs were not included in cost growth estimates, but the DoD is currently trying to include the costs even of sensitive programs wherever possible.

The disadvantages of the hump chart include:

- The cost growth measure includes both historical and anticipated costs, with no separation between them.
- The quantity adjustment is taken directly from the SAR.
- Development and production are considered together.
- Both the base-year and the then-year estimates mix different year dollars.
- Programs at different stages are considered together.

2. MCR's Work

In addition to single-program estimates, MCR also provided estimates of cost growth over time aggregated for all programs. As shown in Table 1, unadjusted cost growth is 59.2 percent, quantity-adjusted cost growth is 26.1 percent, and cost growth adjusted for quantity and escalation is 11.8 percent. This type of estimate suffers from the deficiency of being a diverse population of systems in different stages of development and production.

These estimates are also presented grouped by service and by program phase. Estimates of average cost increases are also presented on a per-program basis, on the

grounds that "smaller dollar value systems grow more (on a percentage basis) than large programs, i.e., the poorer performance of the smaller programs is hidden by the greater cost stability of the larger ones." In fact, this calculation did make a difference—merely adding real cost growth and dividing by the total DE resulted in 11.8-percent real growth, while individual programs averaged 22.7 percent.

Table 1. Total Cost Growth for 96 Systems Listed in the December 1983 SARs

	Billions of Dollars					
Dollar Measure	Development Estimate	Development Estimate Adjusted for Quantity	Other Program Changes	Economic Escalation	Current Estimate	Percentage Growth Adjusted for Quantity
Current	\$471.0	\$594.4	\$130.6	\$24.7	\$749.7	26.1%
Base Year	\$312.2	\$351.6	\$41.6	—	\$393.2	11.8%
Difference		\$242.8			\$356.4	46.8%

Unadjusted Growth = (749.7 – 471.0)/471.0 = 59 2%

Growth Adjusted for Quantity = (749.7 – 594.4)/594.4 = 26.1%

Growth Adjusted for Quantity and Escalation = (393.2 – 351.6)/351.6 = 11.8%

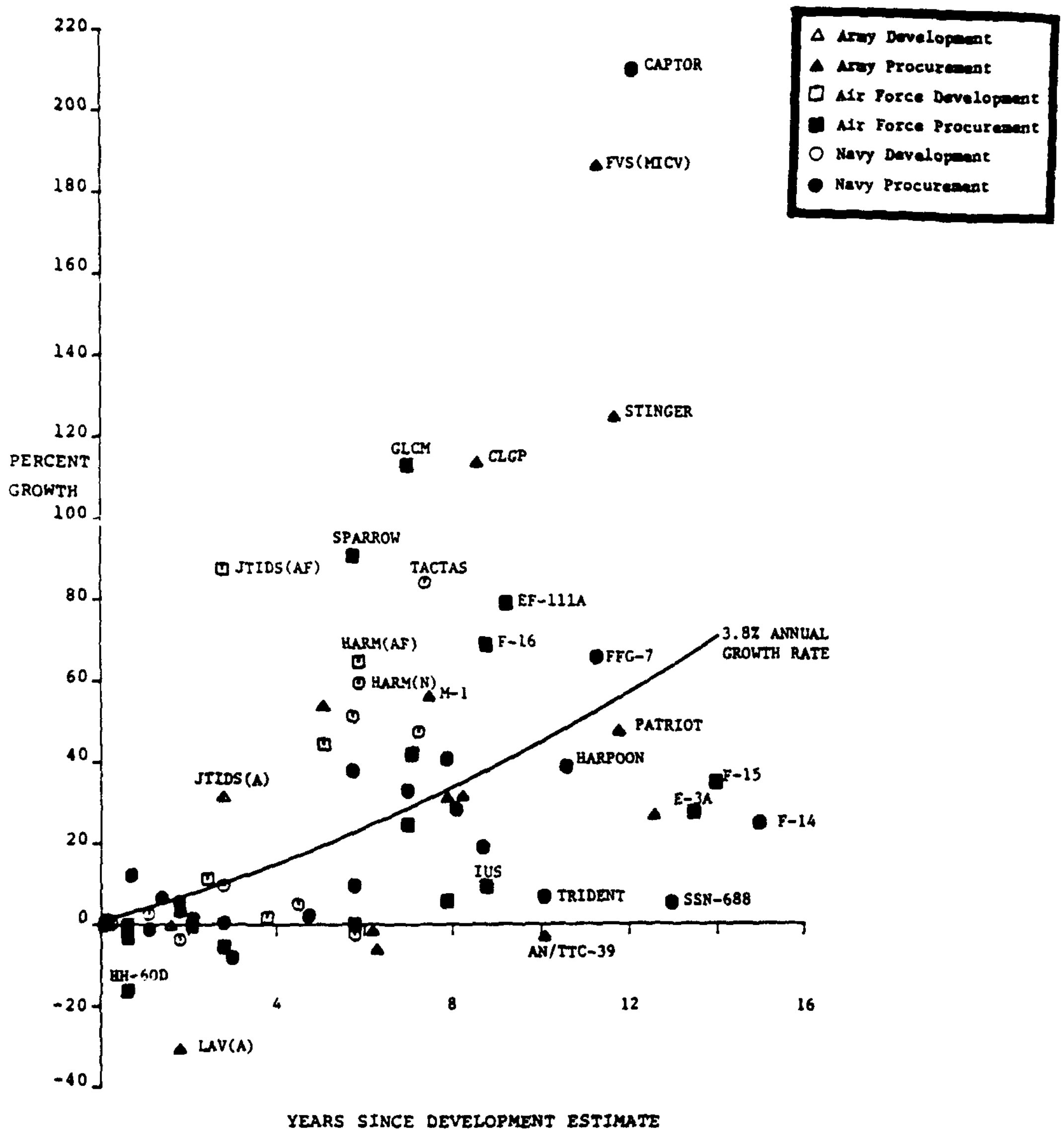
Growth of Inflation = (356.4 – 242.8)/242.8 = 46.8%

Source: Reference [1].

MCR uses these all-program averages to identify individual systems that have growth rates higher than average, and suggests that they be targeted for increased management attention.

They also present an annualized rate of growth per program, by the number of years past the development estimate. Figure 5 shows an annualized growth rate of 3.8 percent [1]. This measure is useful in illustrating that cost growth tends to accumulate as the program ages. This is because information improves and uncertainty lessens as the program ages. It is also true that, because there is an incentive to delay reporting bad news, cost growth tends to accumulate over time.

In later work, analyses were performed that examined growth over four years (1979-1982) of a constant set of 37 systems [3]. This avoids significant problems of a changing population and non-comparable programs.



Source: Reference [1].

Figure 5. Annualized SAR Program Growth Rate as of 31 December 1983

3. RAND's Work

A series of RAND studies on acquisition policy originated with an inquiry by Dr. John S. Foster, then Director of Defense Research and Engineering, about the possibility of comparing the accuracy of program estimates between the 1950s and the 1960s. The first of these studies that indicated a measure of cost growth [6] contrasted outcomes of programs during the 1960s with the experience of programs during the 1950s. Cost, schedule, and performance outcomes were examined. Because of data problems, only undeflated contract dollars were considered. Most cost growth ratios fell between 1.0 and 2.0, but some ranged as high as 5.6. "By means of a statistical model, cost factor numbers were related to time differentials—the time lapse between the construction of a cost estimate and the completion of the program. . . . The analysis showed that, on average, cost estimates for the 1960s were about 25 percent less optimistic than those for programs of the 1950s." (Reference [6, p.vi].)

These cost growth factors suffered from several deficiencies. First, they were derived before there was a process for systematic documentation of program histories. In addition, cost growth for 1950s programs were based on unit costs of production items, while cost growth for the 1960s included both development and production. Although 1950s (and possibly 1960s) cost growth estimates seemed to be based on unit costs, no mention was made of any other correction for changes in quantity. Depending on the magnitude of quantity changes, failure to adjust for the learning relationship could seriously bias the results.

In these estimates, RAND estimated a function:

$$\log \text{CGRatio} = \log a + b\,M$$

where M is the number of months from the initial estimate to IOC.

RAND found that cost growth in the 1960s was lower than that in the 1950s for a given time interval between estimate and delivery. They also found that, as one would expect, the longer the amount of time the estimate covered, the less accurate the estimate was.

The multiple program measures included:

- Average cost growth ratio for a group of programs (programs with Milestone II in the 1970s, for example).
- Average cost growth ratio weighted by program size (which allows for the general pattern of less growth in high-value programs).

- Median cost growth ratio.

- The above measures, excluding programs fewer than three years past Milestone II. This measure makes some adjustment for the fact that cost growth tends to accumulate over time, and it does not make sense to compare mature or completed programs to new programs.

- Average annual linear rate of program cost growth. This is calculated by regressing each program's cost growth ratio by the number of years past Milestone II and forcing the regression through the origin to assume 0 growth at Milestone II. This method again attempts to adjust for the fact that cost growth accumulates over time. The average annual rate was found to be 5 to 6 percent in the 1970s compared with 7 to 8 percent in the 1980s. The method appears to be appropriate for the comparison of programs in the 1960s and the 1970s, when there was a mix of new and mature programs. In the 1980s, however, there was a surge in new programs, and this measure would not be appropriate.

One advantage of RAND's measures was that a quantity adjustment was made. Among the disadvantages were that the quantity adjustment was taken directly from the SAR rather than calculated from a full price-improvement curve, and a biased population of programs, including some very young programs was used.

4. IDA's Work

Asher and Magellet of IDA looked at cost growth over two time intervals, DE to IOC and IOC to latest available estimate [11]. They also looked at average cost growth per year.

The strengths of the Asher-Magellet measure include:

- It separated past from future cost growth.
- It used readily available data and appeared to be replicable.
- It adjusted for inflation.
- It was able to measure development and production separately and together.

The principal weakness of the method was that it lacked a quantity adjustment.

In the study on effective initiatives previously mentioned [9], the aggregation of programs was done more correctly. Programs were excluded from the analysis if they had fewer than three years of experience in either development or production—e.g., programs with fewer than three years of development experience had no measures calculated for

them, and programs with fewer than three years of production experience had only the development measure calculated.

Also in that study, programs were weighted by dollar size when cost growth percentages were compared. Cost growth was analyzed by decade, by equipment type, by new vs. modification program, and by program phase. A separate analysis using only completed programs was also performed.

5. Compound Growth Rate

Biery [12] developed a method designed to correct for the distortions of program age. Biery used annual compound growth rates to make this adjustment:

$$\text{Cost Growth Ratio} = A\,(1 + R)^{Y}$$

where A is forced through a cost growth ratio of 1.0, and

R = the compound growth rate to be estimated

Y = the number of years from the development estimate.

These growth rates were used to compare the records of programs by decades. Biery found that programs of the 1970s had a lower compound annual growth rate than programs of earlier decades, but that the 1970s programs tended to last longer.

6. Time-Series Approaches

Sapp [13] evaluated the cost accounting approach and the cost-category approach, but was most interested in what he calls the predictive-functional approach to measuring cost growth. He views the cost-schedule-performance triad as a response surface. Figure 6 shows a performance contour map. For a given level of performance, cost is a convex function of schedule—e.g., very short programs cost the most, while the benefits to having more time eventually level off. Higher levels of performance require longer development time and/or cost more.

The equations used were adapted from works of such time-series analysts as Box, Wilde, and Draper and Smith. Sapp's concept was

$$Y = f\,(X, B, R),$$

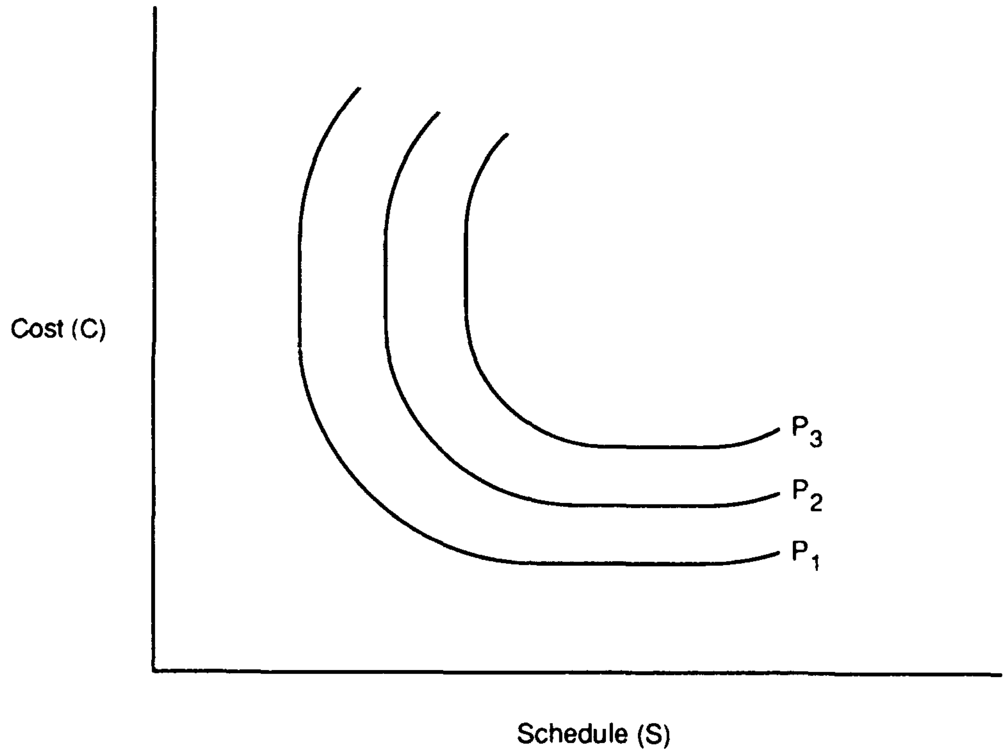
where

Y = cost outcome

X = controllable factors, such as type of contract and contractor

B = boundary factors, such as state-of-the-art technology, the national economy, and inflation

R = random events, such as mobilizations and current events.



Source: Reference [13], Figure 5.

Figure 6. Cost-Schedule Performance Response Surface

Sapp identified the following candidate variables for studying the cost outcome of a modification program:

- Government variables, such as air materiel area, using command, program personnel/program cost, program personnel turnover rate
- Product variables, such as technological advance sought, and kit type
- Procurement variables, such as program priority, funding support, procurement method, number of contracts, type of contract, and time sequencing
- Contractor variables, such as the optimism of proposal assessment, past procurement performance, and percentage of government business.

A study by Launer, Candy, and Carter [14] also used predictive-functional techniques. They found that contract duration is strongly related to cost growth. Inflation

has a small positive effect, and the effect of technology level was significant for Army programs but not for the Navy or the Air Force.

7. Explanatory Models

The IDA study on effective initiatives [9] made a preliminary investigation of the drivers of total program cost growth. Candidates included development outcomes, equipment type, new starts, acquisition initiatives, schedule length, and program stretch. Of these, development schedule growth, program stretch, and development schedule were three significant drivers.

The estimating equation was:

$$\underset{(3.06)}{TPCG = 0.56} + \underset{(2.23)}{0.30} *DSG + \underset{(2.60)}{0.07} * STRETCH + 0.004 * DS$$

$$R2 = 0.38 \qquad F = 9.86 \qquad N = 52$$

where

TPCG = total program cost growth

DSG = development schedule growth

STRETCH = program stretch, the ratio of production schedule growth to production quantity growth

DS = development schedule length.

Numbers in parentheses are t-statistics. The relatively low explanatory power of the equation indicates that further investigation of cost growth drivers would be useful.

In later unpublished work, a start was made toward investigating the impact of contractor and industry conditions on cost growth. Capacity utilization in the aerospace industry was a positive and statistically significant variable in explaining cost growth.

Another IDA study [8] used essentially the same measure to assess the acquisition programs of 19 tactical weapons. This time production cost growth was found to be significantly related to development schedule slip page, but development cost growth was not related to development schedule growth.

Harman [15] added a factor to the model to control for not just the length of development but the level of technical advance. He found that longer, more ambitious programs tended to have higher cost growth. When controlling for technical reach, the 1960s looked worse relative to the 1950s than in a prior study. Cost growth weighted by program size and unweighted yielded similar results.

Harmon, Ward, and Palmer [16] developed schedule estimating relationships for tactical aircraft. Watkins [16] and Norden [17] used similar methods. The development cost equation used could also be applied to track development cost growth. The function used was:

$$Y = K\,(1-e^{-At^{**2}})$$

where

Y = cumulative spending in the given year or other time period

K = eventual cumulative spending

t = months past program start.

That function is intrinsically non-linear and must be estimated using iterative techniques. We would prefer a linear or linearizable function.

Moses [18] has shown that prior to production factors can be identified that are associated with cost overruns or cost savings. He measured cost over(under)runs as the residual of regressions of production cost on measures of technology state-of-the-art and technology advance embodied in aircraft. He found that cost variances are most strongly related to program size, defense spending at the time of program initiation, and the rate of inflation at program start [18, p. 46]:

> Larger programs, because they are more risky or more difficult to manage, tend to be associated with cost overruns. When defense spending is high, cost overruns tend to result. This is consistent with an environment favorable to defense spending leading to acceptance of a higher price by DoD. Cost overruns also tend to follow periods of rapid inflation.

C. NON-DEFENSE STUDIES

1. Energy Process Plants

A series of RAND studies in the late 1970s and early 1980s [19 through 21] considered the problem of cost growth in the civilian world. New energy process plants were found to have problems similar to weapon systems. Costs were consistently underestimated. As with weapon systems, factors that increase costs include scope changes, unanticipated inflation, unanticipated regulatory changes, management practices, bad weather and other miscellaneous factors. Scope changes were regarded as discretionary and were generally not adjusted for.

RAND also analyzed the factors that affected estimation accuracy. A model was developed to help quantify those factors and to evaluate program risk early in the process.

The major factors included: percent new (percentage of the cost representing technology unproven in commercial uses), an index of degree of difficulty posed by process impurities, complexity (number of process steps in the plant), degree of completeness of cost estimate, and project definition (levels of site-specific information and engineering included in the estimate).

2. Growth Accounting

The growth accounting approach is another way of separating controllable and uncontrollable factors in analyzing cost growth.

Edward Denison originated the use of growth accounting in macroeconomics [22]. In his formulation, the factors accounting for economic growth can be expressed as an identity. For example, nonresidential business output is decomposed into total factor input times output per unit of input. The residual is then left as an unexplained, or immeasurable, factor—in Denison's case, technology growth.

Freeland and Schendler [23] used a similar approach in assessing factors accounting for growth in national health expenditures. They were interested in presenting an explanation for why health spending was growing faster than the gross national product (GNP). Over the period 1979-82, for example, real GNP increased less than 1 percent, while real national health expenditures increased nearly 13 percent. This sort of large growth led to initiatives to control costs and improve management of the health care system on the part of Congress, the states, and private industry.

The accounting approach used by Freeland and Schendler (which is used both to account for present spending levels and to predict future spending levels) is:

> National health spending cost growth ratio = Growth in general inflation
> x Growth in health care inflation over and above general inflation
> x Growth in population x Growth in quantity of services per capita
> (measured in terms of hospital days and physician visits, for example)
> x Growth in intensity (the mix and content of services and supplies per
> visit or per day).

The first four factors are measurable, the last, intensity, is the residual. In terms of policy, general inflation and population growth would be viewed as beyond the control of policymakers. However, the rate of health care inflation relative to general inflation might be reduced by policy measures, as could growth in quantity and intensity. This model, then, illuminates the factors relating to health care cost growth.

How is this relevant to acquisition cost growth? The same sorts of factors might be developed for defense policy, as a measure of affordability of programs. At the individual program level, the SARs, as noted above, follow a growth accounting approach. However, the categories as defined do not represent a simple accounting identity. Additionally, a multiplicative model using ratios would be more analytically sound than the additive model used in the SARs.

A strict model at the program level could consist of:

System cost growth ratio = Growth in general inflation x Growth in the appropriate deflator's inflation over and above general inflation x Growth in program quantity per year x Growth in program schedule x Residual unexplained factors.

The problem that needs to be worked out with this model is that the model assumes a linear relationship between program quantity and cost. In major systems, there is a price improvement relationship that needs to be incorporated.

The implementation of this approach in the cost growth literature is imperfect and is complicated by the learning relationship. We would like to be able to look at something like:

Cost Growth over DE = (Quantity Growth) * (Price Growth).

III. NEW MEASURES OF COST GROWTH

There is no single "right" measure of cost growth. Rather, there are families of measures for different purposes. In the previous section, we discussed some existing measures and their strengths and weaknesses. In this section, we elaborate on what could be done to improve existing measures.

A. SINGLE-PROGRAM MEASURES

A cost growth measure, to the extent possible, should be constructed to shed light on the reasons for the cost growth. In the eyes of the cost analyst, cost growth should be adjusted for factors beyond the control of the cost analyst.

The ideal program baseline is one made early enough to allow OSD to monitor the program, yet late enough to be a realistic estimate of a well-defined system. The convention used by OSD has until recently been to measure from the start of full-scale development, when the system is defined.

It is important to remember that, even with an ideal baseline, what we are doing is measuring the outcome of a forecast, which involves both a total cost estimate and a forecasted spending stream. The original forecast may have been unrealistic. There are often perverse incentives in estimating the cost of major programs. These incentives vary depending upon conditions at the time of the estimate.

We believe that it is important to maintain the measurement of cost growth from the development estimate, rather than to use the new production estimate if the purpose of the measure is to evaluate cost analysis and acquisition policy.

Also, since we are working from an OSD point of view, we want to adjust for inflation. If we were working from a congressional point of view, we would not, since Congress budgets in nominal dollars, not real dollars.

In past SARs, program offices were sometimes allowed to develop customized price indexes for inflation adjustments on the grounds that the mix of inputs and the spendout rate for their particular program was unique. This practice concealed a great deal of mischief. If the input price index can be specified in enough detail and the weights customized, it is possible to account for virtually all cost growth in terms of inflation.

OSD now prescribes deflators. However, they are still specific to the service, to the appropriation (RDT&E vs. procurement), and to the equipment type (e.g., Air Force aircraft vs. Navy missiles).

From the standpoint of congressional budgeting, it would be useful to show inflation both in terms of the prescribed OSD deflators and in terms of historical and projected broader measures such as the implicit GNP price deflator. In terms of cost analysis, when we are trying to track program management effects, it is simpler and more convenient to use base-year dollars throughout the calculations. However, to the extent to which inflation is underpredicted by OSD, programs will tend to exhibit then-year dollar cost growth even though the program achieved its goals in real terms. This is because escalation rates used in developing the out-year budgets are likely to be different from the rates used for the development estimate.

In this report, we do not treat the issue of inflation completely, but instead work in base-year dollars throughout. Nevertheless, this is an important area for future study.

We further recommend that only two adjustments be made for program change—inflation and quantity. They should be performed in that sequence. From the point of view of the cost analyst, the inflation adjustment can be avoided by always working in constant dollars.

1. Development Cost Growth

Total cost growth in development is relatively easy to measure. It is simply the ratio of the development estimate to the current estimate. No quantity adjustment is necessary, because the purpose of development articles is to get the system to IOC. In any event, the number of development articles purchased is usually not substantially different from the number planned.

The one judgment call in measuring development cost growth is whether or not to cut off development spending at IOC. Often, there is real cost growth past IOC. However, substantial expenses in the RDT&E budget after IOC are typically for modification programs, which represent a clear scope change from the baseline DE program.

Quantity is generally not adjusted for in development programs. This is because the bulk of cost in a development program is fixed rather than variable with the number of development articles.

2. Production

Cost growth in production is complicated by changes in quantity procured. A typical solution when funding is scarce is to "stretch" the program—to buy fewer units over the same period of time or to buy the planned number of units over a longer period of time.

Program stretch can greatly increase cost estimates. This is because defense contractors have both fixed and variable costs. Variable costs change when quantity changes. However, fixed costs vary not with respect to quantity but with respect to time. Even if the same number of items are being acquired, costs might be higher because of fixed costs incurred when the program is stretched.

There shou'd be an appropriate quantity adjustment using the current estimate actual price-improvement curve and adjusting back to the development estimate quantity. This adjustment should be done using the price-improvement curve on actual data, rather than by using the quantity adjustment in the SAR, which represents only a best estimate and is based on the DE curve, not the CE curve.

3. Separating Past From Future

A major defect of past measures of cost growth is that they essentially compare a projection to a projection. Thus, the total program cost growth measure for an ongoing program included both cost growth that already has occurred and an estimate of future cost growth. It would be useful to separate these two categories, since the former is water under the bridge, while the latter may still be prevented. This is particularly important during the development process, as programs last longer.

4. Accounting for Cost Growth

A growth accounting approach that measures only those variables that can be measured with precision, and leaves the rest as a residual, could be a useful policymaking tool. As indicated above, this type of analysis can be powerful, but it is hampered by the problem of adjusting for learning and rate. Nevertheless, if these problems could be resolved, the DoD would have a tool for separating controllable and uncontrollable factors in cost growth. At the very least, one would want to separate inflation as measured by the GNP deflator from defense-related inflation.

5. Assessing Program Risk

One can also assess program risk by looking at more sophisticated measures for specific programs. For example, total program cost growth is related to development schedule growth, program stretch, and development schedule length [9]. These factors can be used to predict cost growth early in a program.

B. MULTIPLE-PROGRAM MEASURES

1. Tracking Cost Growth

The measure proposed by the Congressional Budget Office (CBO) and implemented by OSD is designed to give an indication of the DoD's progress in curtailing cost growth in weapons acquisition through various management initiatives. As noted by CBO [7, p. 7]:

> Not all program-cost changes are the responsibility of DoD management. Many factors influencing estimated costs are beyond the control of the program manager. For example, the unexpected capability of a potential enemy to jam the guidance system of an air-to-air missile may require an engineering change to counter it. Therefore, program-cost changes, excluding economic and quantity changes, can serve only as a very general indication of the impact of the Department's efforts to improve acquisition programs.

We recommend the following:

- The three types of baselines should be reported separately. Currently, all three types of baselines are mixed.
- Percentage cost changes should be calculated separately for each program and then weighted by program size in common-year dollars. This eliminates the problem, noted by CBO, that the base year varies by program, so even constant-dollar data do not have a common base year.
- Development and production cost growth should be calculated separately. The stage of completion should be noted, so analysts could exclude production cost growth calculations for programs not yet in production, if desired.
- Annual growth rates, should be calculated, and programs should be grouped by percentage of planned or actual schedule completed.
- Programs should be compared with past averages by equipment type. Tyson, Nelson, Om, and Palmer [9] provide averages for comparisons of this type.

2. Cost Growth Estimating Relationships

In the same way that they use cost estimating relationships, cost analysts need to use cost growth estimating relationships. These estimating relationships would allow cost

analysts to assess program risk or to assess the potential outcome of a partially completed program in a more rigorous way.

In the studies attempting to evaluate the success of policies in a particular era, programs are typically grouped by the year of full-scale development. This is a useful classification, because the acquisition strategy is typically decided at this time. However, programs vary in terms of length, and policies have changed unusually rapidly in recent years. It would be useful to have a measure that took into account for each program the specific combination of policies used in each program. This would allow for the evaluation of acquisition policy in a more precise way than the approach of dividing programs into decades or half-decades allows.

Among the factors that need to be considered are:

- Variables beyond the control of the individual program, such as:
  - The overall tightness of the defense budget at the time of the cost estimate, which is a proxy for the incentive to estimate costs on the low side. If the budget is very tight, there may be pressure to be optimistic about costs in order to sell the program. This variable could be measured by the real percentage change in the defense budget, or the defense acquisition budget, using a five-year moving average.
  - Inflation in the economy. This variable could be measured by the percentage change in the GNP deflator.
  - Defense inflation. This can be measured by the ratio of changes in the defense deflator relative to the GNP deflator.
  - Market conditions in the industry, as measured by capacity utilization. This is a proxy for the "hunger" of the industry to complete the program in a cost-effective manner. If capacity utilization is relatively low, then contractors are likely to be willing to reduce costs in order to keep business.
  - Variables relating to general economic conditions, including GNP growth.
- Program characteristics:
  - Acquisition initiatives
  - Contract types
  - Technology reach
  - Planned schedule length, relative to averages for past programs
  - Program size, as a measure of managerial challenge to the contractor and the government

- Equipment type
- Program stretch. (Ideally, this would be treated as an endogenous variable. Some stretches occur largely because of general constraints on the budget; others occur because the program itself is in trouble.)

- Contractor variables:
  - Contractor dummy (It would be useful for the government to find out whether or not some contractors perform consistently better than others.)
  - Percentage of defense to total business base
  - Financial ratios showing capital intensity (e.g., capital/labor or capital/assets ratio) and profitability
  - The percentage subcontracted
  - Contractor size relative to program size
  - Program size relative to total business base.
- Random events such as mobilizations and threat changes.

These models could be used in a number of ways. They can be used to explain differences in past behavior or to evaluate the impact of policy in different acquisition eras. They could also be used to assess program risk—e.g., predict cost growth—at the beginning of a program, as well as to simulate the impact of changes in the program. OSD could also use these models to assess the realism of submissions by the military services. In an era in which budgeting for affordability is particularly important, these projections would be useful.

REFERENCES

[1] McNichols, Gerald R., James P. McFarland, Bruce J. McKinney, and Stephen J. Balut. "The Problem of Cost Growth (Analysis of Weapon System Costs as of December 1983)." TR-8319-2, Management Consulting and Research, 30 April 1984.

[2] Dews, Edmund, Giles K. Smith, Allen Barbour, Elwyn Harris, and Michael Hesse. "Acquisition Policy Effectiveness. Department of Defense Experience in the 1970s." R-2516-DR&E, The RAND Corporation, October 1979.

[3] McNichols, Gerald R. "Analysis of Defense Cost Overruns." Management Consulting and Research, Briefing to Washington OR/MS Council, 15 February 1989.

[4] McNichols, Gerald R., and Bruce J. McKinney. "Analysis of DoD Weapon System Cost Growth Using Selected Acquisition Reports (as of 31 December 1980)." Presentation to the Sixteenth Annual DoD Cost Analysis Symposium, October 1981.

[5] Alvin J. Harman. "Acquisition Cost Experience and Predictability." Paper P-4505, The RAND Corporation, January 1971.

[6] Perry, R. L., D. DiSalvo, G. R. Hall, A. J. Harman, G. S. Levenson, G. K. Smith, and J. P. Stucker. "System Acquisition Experience." RM-6072-PR, The RAND Corporation, November 1969.

[7] U.S. Congress, Congressional Budget Office. "A Review of the Department of Defense December 31, 1983 Selected Acquisition Report (SAR)." Special Study, July 1984.

[8] Gogerty, David C., Bruce M. Miller, J. Richard Nelson, and Paul R. Palmer, Jr. "Acquisition of Contemporary Tactical Munitions." 2 Volumes, Paper P-2173, Institute for Defense Analyses, March 1990.

[9] Tyson, Karen W., J. Richard Nelson, Neang I. Om, and Paul R. Palmer, Jr. "Acquiring Major Systems: Cost and Schedule Outcomes and Acquisition Initiative Effectiveness." Paper P-2201, Institute for Defense Analyses, March 1989.

[10] U.S. Department of Defense, OASD(C). "Preparation and Review of Selected Acquisition Reports." DoD Instruction 7000.3-G, May 1980.

[11] Asher, Norman J., and Theodore F. Magellet. "Estimating the Cost Growth of Weapon Systems." Paper P-1494, Institute for Defense Analyses, June 1980 (Revised September 1984).

[12] Biery, Frederick P. "The Accuracy of Military Cost and Schedule Forecasts." *Journal of Cost Analysis* (Spring 1986).

[13] Sapp, Richard S. "The Procurement Process and Program Cost Outcomes: A Systems Approach." Air Force Institute of Technology.

[14] Launer, Robert L., Harold F. Candy, and Shirley H. Carter. "Cost Growth—Effects of Contract Size, Duration, Inflation, and Technology." Report PRO-007-2, U.S. Army Procurement Research Office, May 1972.

[15] Harmon, Bruce A., and Lisa M. Ward. "Assessing Acquisition Schedules for Tactical Aircraft." Paper P-2105, Institute for Defense Analyses, February 1989.

[16] Watkins, Harry. "An Application of Rayleigh Curve Theory to Contract Cost Estimation and Control." M. S. Thesis, Naval Postgraduate School, March 1982.

[17] Norden, Peter V. "Useful Tools for Project Management." In M. K. Starr, ed., *Management of Production.* Middlesex, England: Penguin, 1970.

[18] Moses, O. Douglas. "Estimating and Explaining the Production Cost of High-Technology Systems: The Case of Military Aircraft." NPS-54-89-07, Naval Postgraduate School, May 1989.

[19] Merrow, Edward W., Stephen W. Chapel, and Christopher Worthing. *A Review of Cost Estimation in New Technologies: Implications for Energy Process Plants.* R-2481-DOE, The RAND Corporation, July 1979.

[20] Merrow, Edward W., Kenneth E. Phillips, and Christopher W. Myers. *Understanding Cost Growth and Performance Shortfalls in Pioneer Process Plants.* R-2569-DOE, The RAND Corporation, September 1981.

[20] Merrow, Edward W. "Cost Growth in New Process Facilities." P-6869, The RAND Corporation, April 1983.

[21] Denison, Edward F. *Accounting for United States Economic Growth, 1929-1969.* Washington, D.C.: The Brookings Institution, 1974.

[22] Freeland, Mark S., and Carol E. Schendler. "Health spending in the 1980s: Integration of clinical practice patterns with management." *Health Care Financing Review,* Volume 5, No. 3. (Spring 1984), pp. 1-68.